Sequential Query Expansion using Concept Graph

Saeid Balaneshin-kordan and Alexander Kotov
Computer Science Department, Wayne State University

Objectives

- To discover **effective latent concepts** for query expansion from from manually and automatically constructed concept graphs (or semantic networks), in which the nodes correspond to words or phrases and the typed edges designate semantic relationships between words and phrases.
- To leverage **large and dense concept graphs**, in which the number of candidate concepts that are related to query terms and phrases and need to be examined increases exponentially with the distance from the original query concepts.

Optimization Problem

- **Objective Function:** total number of **evaluated** concepts
- **Constraint:** precision of retrieval results

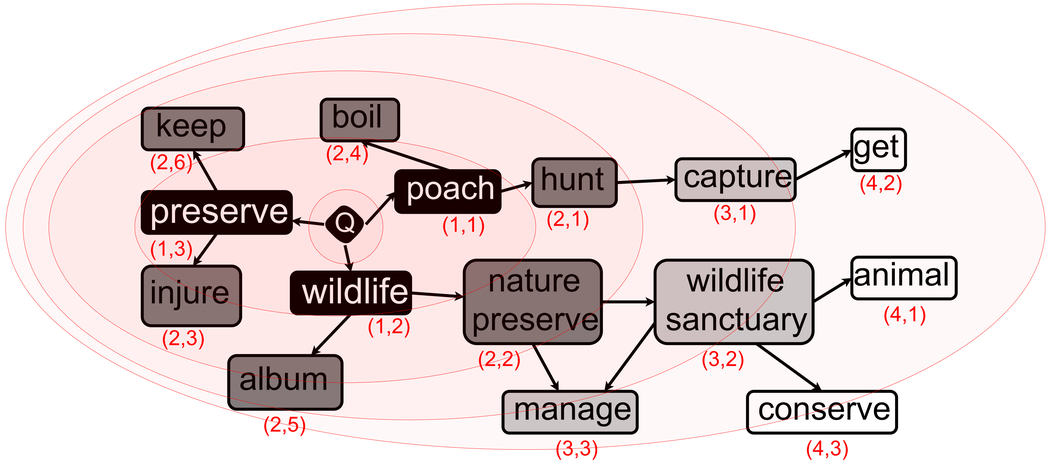
$$\min_{\tilde{\mathbb{C}}_k^{ut}} \left\{ \sum_{i=1}^{k} N_i \right\}$$
$$\text{s.t.} \quad E(\tilde{\mathbb{R}}_\Lambda; \mathbb{T}) > \theta_Q \ ,$$

Approximate Solution

1. The candidate concepts are **sorted** using computationally **inexpensive features**.
2. This sorting is utilized in the **second stage** to sequentially **select expansion concepts** by using computationally **expensive features**.

Criterion	Decision
If $\tilde{Q}_r(C_{(i,j)}) \geq \beta_U$	Select concept $C_{(i,j)}$ & Continue with the same concept layer
If $\beta_L \leq \tilde{Q}_r(C_{(i,j)}) < \beta_U$	Discard concept $C_{(i,j)}$ & Continue with the same concept layer
If $\tilde{Q}_r(C_{(i,j)}) < \beta_L$	Discard concept $C_{(i,j)}$ & Move to the next concept layer

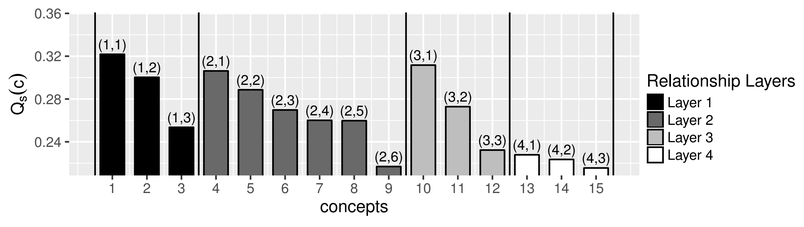
Concept Graph for Query: poach wildlife preserve



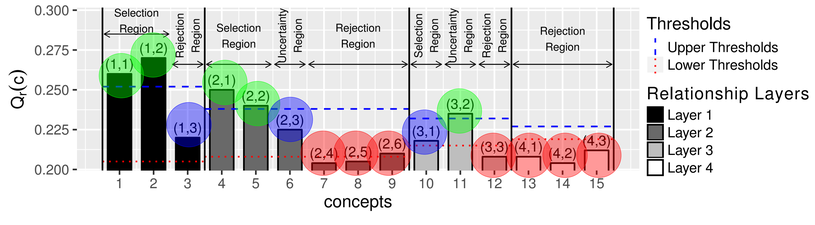
- Concept Graph: ConceptNet 5
- The 1st number in parentheses: concept layer, the 2nd number: index of a concept in the concept layer.

Proposed Sequential Concept Selection

Stage I:
Initial Concept Sorting



Stage II:
- Sort Concepts at Each Layer according to $\tilde{Q}_s(C_{(i,j)})$
- Sequential Concept Selection



Decision	Criterion
Select concept & continue	$\tilde{Q}_r(C_{(i,j)}) \geq \beta_U$
Discard concept & continue	$\beta_L \leq \tilde{Q}_r(C_{(i,j)}) < \beta_U$
Discard concept & move to next layer	$\tilde{Q}_r(C_{(i,j)}) < \beta_L$

Actions

Select
Continue
Discard
STOP

Number of Evaluated Concepts: 11 (26% less)

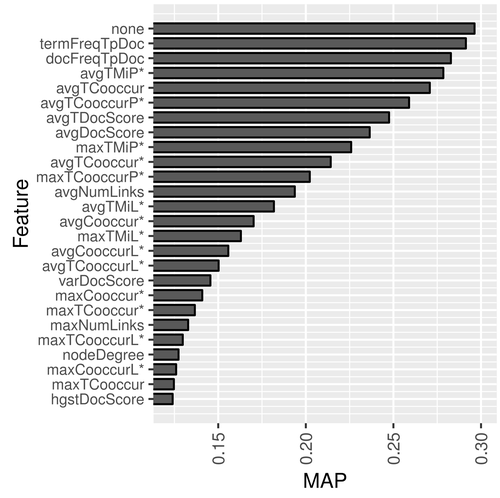


Figure 1: Feature Ablation

Conclusion

The main contribution of this work is a two-stage method for sequential selection of effective concepts for query expansion from the concept graph. The proposed method is formulated as an optimization problem with the goal of evaluating the least possible number of candidate concepts needed to ensure a given precision of retrieval results.

References

Contact Information

- Web: http://www.balaneshin.com/
- Email: saeid@wayne.edu
- Phone: +1 (313) 316 6723

WAYNE STATE

CIKM 2016 Indianapolis
October 24-28, 2016